POLAR PAM 42 POLARPAM

APR 1 1986

Reprinted from CHEMICAL ENGINEERING PROGRESS, January, 1986

Topical Databases: Cold Regions Technology On-Line

One of the key advantages of a topical database over a discipline-based one is the specialized thesaurus used for its subject and indexing terms, which enables records to be accessed more quickly. In addition to saving time, it also cuts down on hourly connect charges.

Nancy C. Liston, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, N.H. 03755
Marilee E. Winiarski, SDC Information Services, Santa Monica, Calif. 90406

A fundamental, and often difficult, question facing all researchers is where to begin an information search. Both hard copy and electronic information sources can be classified in many ways, but it is most practical to distinguish between general, discipline-based resources—*Chemical Abstracts* (for chemistry) or *COMPENDEX* (for engineering)—and the special interest, topical database. An excellent example of the latter is the *Bibliography of Cold Regions Science and Technology*, known as the COLD database on-line. Like other topical resources, COLD resembles its larger cousins in format and search formulation. However, COLD provides the specialist a unique information tool that was developed by cold regions experts across the disciplines to answer a defined range of questions in great depth.

The Library of Congress has been preparing the *Bibliography of Cold Regions Science and Technology* annually since 1951, under the sponsorship of the U.S. Army Corps of Engineers Cold Regions Research and Engineering Laboratory (CRREL), Hanover, N.H. This laboratory has a special mission to understand the characteristics of the cold regions of the world and to apply this knowledge to enable people to live and work in a cold environment. The *Bibliography* has been compiled to enhance research efforts in the areas of snow, ice, frozen ground, climatology, military operations, navigation on ice, ice engineering, construction, environmental, chemical, and civil engineering in cold regions, and the operation of materials and equipment in cold temperatures.

The CRREL originally compiled the *Bibliography* on individual 3 × 5 in. cards available either for filing in a standard catalog or in a bound book volume. Items were indexed by subject and author, and abstracted. Since CRREL eliminated the card format and discontinued its abstracting operation in 1968, the Library of Congress has been preparing the *Bibliography* on MARC II tapes.

The hard copy edition of the *Bibliography* continues to be prepared monthly and cumulated annually into bound volumes. The laboratory is just completing Volume 39 for 1985. These bound volumes are available for sale from the National Technical Information Service, Springfield, Va. 22161. Each volume is compiled in two sections: the numerically arranged full citation section and an author/subject index section. For historical research, three five-year cumulative author/subject indexes have been compiled. The first 22 volumes, accumulated in one author/subject index, are available from the National Technical Information Service.

The COLD database

As the field of information science turned more and more to computerized literature searching, the Cold Regions Laboratory initiated an exclusive agreement with System Development Corp. to mount a computerized version of the *Bibliography* on its information retrieval service, ORBIT. This move was

JUL - 8 1986
No.:
Free

made not only to enhance researchers' ability to locate needed information through the advanced retrieval capabilities of the computer, but to insure international access to the *Bibliography*. The ORBIT system also contains complementary files such as Engineering Index (COMPENDEX), Engineering Meetings (EIMET), National Technical Information Service (NTIS), Physics Abstracts (INSPEC), and GeoRef. In 1977, the COLD database was made available to the public.

In the on-line database, the *Bibliography of Cold Regions Science and Technology* is merged with a complementary resource, the *Antarctic Bibliography*. This bibliography, also prepared by the Library of Congress, is sponsored by the National Science Foundation. The emphasis in the *Antarctic Bibliography* is more general, encompassing materials on expeditions, equipment and supplies, and political geography, as well as the biological and geological sciences, chemistry and engineering. Coverage of the *Antarctic Bibliography* dates back to 1962 in the database. The *Bibliography of Cold Regions* is covered from 1968 to the present.

As of March, 1985, the COLD database contained approximately 98,000 records, with updates adding 1,400 records quarterly. The majority of records from the *Antarctic Bibliography* contain abstracts. However, abstracts are included for only approximately 20% of the citations from the *Bibliography of Cold Regions*. The materials cited from both sources include maps, dissertations, patents, journal articles, conference proceedings, government reports, and monographs. COLD includes all Cold Regions Research and Engineering Laboratory reports, as well as the reports of its predecessors, the Snow, Ice and Permafrost Research Establishment and the Arctic Construction and Frost Effects Laboratory.

The database is international in scope. In the *COLD Regions Bibliography* alone, approximately 60% of the cited references describe Russian language material, 20% other foreign language materials, and the remainder English language documents. Japanese and Canadian researchers are also heavy contributors to the file. All items cited are published materials rather than ongoing research plans. Anyone interested in ongoing unpublished research on cold regions topics should consult other information resources describing government research in progress.

Searching the COLD database

Citations in the COLD database include two types of information: fields describing the documents themselves and fields describing the subject content of the documents. The subject-related fields are of primary interest to most searchers. Unlike the searcher using hard copy, who is limited to precoordinated terminology from a subject index, the on-line searcher may search words from the Index Term, Title, and Abstract fields in infinite combination. Using Boolean logic (and, or, not) and proximity operators, which specify word order and adjacency, the searcher can reach the established search goals.

For example, a searcher may choose to maximize recall, i.e., the number of relevant documents retrieved in a literature search. In this case, it is necessary to identify keywords and possible synonyms for each main concept of the search. If the search topic in the COLD database is "implications of highway design, construction, and maintenance on snow removal and snow removal equipment," the concept of highway might be searched by ORing together these terms: highway, superhighway, expressway, freeway, toll road, parkway. This search statement would retrieve all the documents containing any of these synonyms. This set would then be ANDed together with two other sets, one representing design/construction /maintenance concepts and the other snow retrieval terminology. The final retrieved set would contain only documents with all three concepts. Terms would be searched in the Basic Index, a grouping of all terms in the subject-related fields. By using the Basic Index, the searcher avoids the need to know if a word

As of March, 1985, the COLD database contained approximately 98,000 records, with updates adding 1,400 records quarterly. The majority of records from the *Antarctic Bibliography* contain abstracts. However, abstracts are included for only approximately 20% of the citations from the *Bibliography of Cold Regions*.

Although not limited to index terms, the experienced on-line searcher will make extensive use of the subject indexing created by the database supplier. Every database uses a different thesaurus or authority list for its subject and indexing terms, one usually created or modified for its particular audience.

is mentioned only in the Title field or the Abstract, etc.

If, on the other hand, the searcher wishes to maximize the percentage of documents retrieved that are relevant to the search question (precision), a different, more specific search strategy would be necessary. The searcher would specify the exact relationships between words—e.g., highway or road not only in the same document as construction, but immediately preceding the latter term: (highway or road) (w) construction. Additionally, the searcher could limit the search terms to a particular field, such as Title or Index Term, in the belief that words appearing in these more structured fields are critical to the content of the document.

Although not limited to index terms, the experienced on-line searcher will make extensive use of the subject indexing created by the database supplier. Every database uses a different thesaurus or authority list for its subject and indexing terms, one usually created or modified for its particular audience. One of the key advantages of a topical database is this specialized vocabulary, which allows the searcher to increase the precision of the search formulation. For example, the record "On the Size Distribution of Antarctic Icebergs" will be retrieved in COLD by any of the following index terms: iceberg—distribution; iceberg—ablation; ice models; Antarctica; icebergs; distribution; models (COLD, accession number 34002446). The searcher looking for documents on the distribution of icebergs can precisely describe his search with a single index term. In COMPENDEX, however, the same document is indexed more generally: mathematical models; probability; oceanography; icebergs (COMPENDEX, accession number 80-068587). The searcher looking for iceberg distributions will eventually find this document, but probably with a broad search that also retrieves less relevant documents—at the cost of time and increased connect hour charges. Geographical indexing in COLD is also more detailed than in many general databases, since COLD users are more likely to be searching for information on specific Antarctic and Arctic locations. Records indexed to "Antarctic" in general databases will be indexed to the more specific level, "Antarctica—Lassiter Coast" or "Antarctic—Arctowski Station," in COLD.

As these last examples indicate, it is critical that the searcher of topical databases knows how the database vocabulary is controlled, both in terms of depth and technical content, and in terms of style: hyphenation, punctuation, word inversion, and use of compound phrases. The best source of CRREL vocabulary at present is the latest volume of the *Bibliography of Cold Regions*. With experience using the database and reference to the printed *Bibliography*, the searcher will find COLD's controlled vocabulary a powerful searching tool.

Although most database searchers rely on subject-related fields when initially formulating their search strategies, the ability to limit a search strategy by year, language, or type of document is also important. By restricting certain bibliographic elements, the searcher avoids retrieving citations, which, although pertinent to the search terms, are not useful due to age, lack of availability, language, etc. The bibliographic fields available in the COLD database include the Author, Source (journal name, report numbers, etc.), Organizational Source, Document Type, Language, File Segment, Issue Number (*Bibliography of Cold Regions* only), Publication Year, and Accession Number. A researcher may directly search any of these elements independently or in combination with subject-related search terms.

In terms of vocabulary and search formulation, the searcher will approach specialized databases like COLD somewhat differently than general databases. However, the rules to ensure a high quality search are essentially the same for all databases. If a search intermediary is employed, it is preferable that the person who conducted the original reference interview actually perform the search, analyze the results, and return them to the requester for comment. This procedure guarantees continuity and, if there is an error, makes it easier to discover and correct. The information seeker should carefully and comprehensively de-

BOREAL INSTITUTE
LIBRARY

fine terms and identify any limitations, such as date or language, to be imposed on the retrieved set. If possible, a few relevant references should be located, as using a known reference can lead the searcher to appropriate index terms and additional synonyms.

The searcher should use the wide range of retrieval system commands that make computerized literature searching such a powerful tool. The techniques used to make on-line searching efficient and cost-effective include truncation of root terms to pick up multiple variations of words, search statement review and editing, and transfer of search terms from print records to search statements without rekeying. In addition, the searcher should take advantage of the interactive nature of on-line searching to revise and refine the searching strategy to obtain the desired result.

Once a strategy is perfected, the searcher may wish to consider storing the search on-line. The search can then be re-executed each time the database is updated, thus providing the researcher with the latest information on a topic. On the COLD database, automatic execution of stored strategies in database updates, also called Selective Dissemination of Information, is available. Since cold regions research is multidisciplinary, searchers may develop stored searches that can be executed regularly in a variety of related databases, including the World Patents Index, the Metals Data File, STANDARDS and CORROSION. Although the searcher may find some duplication, the source materials covered by these databases differ widely. Therefore, a search defined and developed in a topical database like COLD can be profitably expanded into other topical and general files.

Obtaining indexed documents

Of course, every citation in the COLD database contains enough bibliographic detail to identify the original document to a library, document supplier, or the CRREL. Citations retrieved in a search of the COLD database may be printed at the searcher's terminal for immediate use. Or the searcher may choose to take advantage of cost reductions and report quality laser printing by printing off-line at the vendor's facility. In either case, the searcher can request as much or as little detail as necessary, from the complete record to any single field. Once the retrieved citations are printed, the researcher can determine which references will best fill the information need.

Hard copy retrieval of documents cited in on-line databases can be a problem, especially for researchers in small engineering or industrial firms with limited library resources. A unique feature of the *Bibliography of Cold Regions* is the availability on microfiche in the CRREL library of every item generated from a literature search. Like many topical database suppliers, CRREL indexes a substantial number of rather obscure documents. The CRREL microfilm library increases the value of the COLD database immensely by providing easy and timely access to documents which might otherwise be unobtainable. The CRREL library provides access to or copies of items retrieved to a limited extent within the confines of U.S. copyright law and the limited staff of the CRREL library.

In conclusion

The usefulness of a topical database such as COLD is ultimately determined by its users. Used independently and in conjunction with other information resources available both in hard copy and on-line, such databases can enhance the efficiency of scientific and technical research. By serving special information needs worldwide in the areas of Arctic, Antarctic and cold research, the COLD database and its related print products continue to assist the researcher concerned with cold regions technology.

Liston

Winiarski

99999 Pam:025.4.036
LIS

LISTON, Nancy C. and WINIARSKI, Marilee E.
Topical databases

Borrower's Name	Date Due

99999 Pam:025.4.036
LIS

LISTON, Nancy C. and WINIARSKI, Marilee E.
Topical databases

Boreal Institute for Northern Studies Library
CW 401 Bio Sci Bldg
The University of Alberta
Edmonton, AB Canada T6G 2E9

University of Alberta Library
0 1620 0327 9526